Cloning and expression of the dihydrofolate reductase-thymidylate synthase gene from *Trypanosoma cruzi*

Pedro Reche[a], Rosalia Arrebola[a], Asunción Olmo[a], Daniel V. Santi[b], Dolores Gonzalez-Pacanowska[a], Luis M. Ruiz-Perez[a,*]

[a]*Instituto de Parasitología y Biomedicina "Lopez-Neyra", Consejo Superior de Investigaciones Científicas, C/ Ventanilla 11, 18001-Granada, Spain;* [b]*Departments of Biochemistry and Biophysics and of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143, USA*

Received 25 October 1993; accepted 3 March 1994

Abstract

We have cloned, sequenced and expressed the *Trypanosoma cruzi* gene encoding the bifunctional protein dihydrofolate reductase-thymidylate synthase (DHFR-TS). The strategy followed for the isolation of positive clones from a genomic library was based on the construction of a probe by the amplification of highly conserved sequences of the TS domain by the polymerase chain reaction. Translation of the open reading frame of 1563 bp yields a polypeptide of 521 amino acids with a molecular mass of 58829 Da. For heterologous expression of *T. cruzi* DHFR-TS in *Escherichia coli*, the entire coding sequence was amplified by polymerase chain reaction and cloned into the plasmid vector pKK223.3. The presence of catalytically active DHFR-TS was demonstrated by complementation of the Thy^- *E. coli* strain χ2913 and the $DHFR^-$ Thy^- *E. coli* strain PA414. The gene is expressed as an active protein which constitutes approximately 2% of the total cell soluble protein. Recombinant bifunctional enzyme and the DHFR domain have been purified by methotrexate-Sepharose chromatography to yield 1–2 mg of active DHFR-TS per litre of culture. Southern and electrophoretic analyses using the coding sequence as probe indicated that the *T. cruzi* enzyme is encoded by a single copy gene which maps to two bands of approximately 990 kb and 1047 kb. It appears that *T. cruzi* is diploid for the DHFR-TS gene which is located on two different-sized homologous chromosomes.

Key words: Trypanosoma cruzi; Dihydrofolate reductase-thymidylate synthase; Protozoal enzyme; Heterologous expression; Folate metabolism

*Corresponding author. Tel: 34-58-203802; Fax: 34-58-203323; e-mail: LMPerez@samba.cnb.uam.es.

Note: Nucleotide sequence data reported in this paper have been submitted to the Genbank™ data base with the accession number L22484.

Abbreviations: DHFR-TS, dihydrofolate reductase-thymidylate synthase; CHEF, clamped homogeneous electric field electrophoresis.

1. Introduction

Dihydrofolate reductase (DHFR; EC 1.5.1.3) and thymidylate synthase (TS; EC 2.1.1.45) catalyze consecutive reactions in the de novo synthesis of dTMP. In most organisms these two enzymes exist as separate monofunctional proteins [1]. In contrast, in protozoa both enzyme activities coex-

ist on the same 55–70-kDa polypeptide chain, with the DHFR domain at the amino terminus and TS at the carboxy terminus [2,3]. TS and DHFR have received a considerable amount of attention and have been extensively studied as potential chemotherapeutic targets.

Parasitic protozoa cause major world health problems, and one approach towards the development of chemotherapeutic agents involves the identification and exploitation of unique differences between the molecular structure of parasitic enzymes and their mammalian counterparts. DHFR inhibitors such as trimethoprim and pyrimethamine have been widely used as antimalarial agents but classical antifolates have not been clinically useful in the treatment of diseases caused by other parasitic protozoa. The fact that DHFRs from different species show profound differences in their primary sequences is a clear example of the divergence throughout evolution of this enzyme. Such differences should allow for the development of species-specific drugs. In this regard, a series of 2,4-diamino-5-substituted benzyl pyrimidines have proven to be effective as inhibitors of purified *Leishmania major* DHFR-TS [4]. These pyrimidine analogs are more potent inhibitors of *Leishmania* promastigote and amastigote growth than the classical anti-microbial antifolates and represent possible lead compounds in the future development of new antiprotozoan agents.

High level expression systems for *Leishmania* DHFR-TS have been developed using heterologous expression systems [5]. Likewise an expression system for the *Plasmodium falciparum* bifunctional enzyme has been obtained in *Escherichia coli*, although the enzyme is poorly expressed [6]. The availability of several protozoan DHFR-TSs has important implications for future drug design. It is reasonable to assume that a sufficient knowledge about the structure of DHFR-TSs from different protozoa and the characterization of their interaction with inhibitors would open routes to the design of specifically targeted drugs.

At present a crystal structure of the bifunctional protein from protozoa is not available. Comparison with the resolved structures of the monofunctional counterparts from *E. coli* and *Lactobacillus casei* has allowed for predictions on structural domains and the assignment of specific roles to residues involved in substrate binding and catalysis. In any case, the obtention of adequate amounts of enzymatic protein is required for kinetic and structural studies that might lead to the development of selective drugs against protozoan DHFR-TS.

2. Materials and Methods

Reagents. The *E. coli* K12 strain χ2913 (thy A 572), which carries a deletion in the TS gene, was a gift from Dr. R. Thompson (University of Glasgow), and the TS and DHFR-deficient *E. coli* strain PA414, was a gift from Drs. Ahrweiler and Frieden (University Medical School, St Louis). The *T. cruzi* Y strain was used for genomic DNA and RNA preparations. Other reagents have been previously described [5,7] or were of the highest purity commercially available. Oligonucleotides ON1 (5′GGAATTCCRTNTAYGGNTTYCARTGG3′), ON2 (5′GGGATCCCGCNATRTTRAANGGNAC3′), ON3 (5′TTGAATTCATGTCGCTGTTTAA3′) and ON4 (5′GTAAGCTTCTAAACCGCCAT3′), where R indicates a purine and Y indicates a pyrimidine, were synthesized at the University of California, San Francisco, Biomolecular Resource Center.

DNA manipulations procedure. Total DNA from *T. cruzi* (Y strain) was prepared by phenol extraction [8]. Agarose gel electrophoresis, staining, blotting and hybridization were performed by standard procedures [9]. The hybridization probe for screening for the *T. cruzi* DHFR-TS gene was obtained by the polymerase chain reaction (PCR) technique. The reaction mixture (50 µl) contained 25 pmol of each of the two oligonucleotide primer mixtures ON1 and ON2, 500 ng of *T. cruzi* genomic DNA, dNTPs each at 400 µM, 60 mM KCl, 25 mM Tris HCl (pH 8), 10 mM $MgCl_2$, and 0.1% bovine serum albumin. Amplification was initiated with 1.5 U of Taq polymerase. PCR parameters were 35 thermal cycles consisting of a 30-s denaturation at 94°C followed by a 21-s annealing period at 45°C and a 1-min extension period at 70°C. Amplification products were purified on a 2%

agarose gel, digested with *Eco*RI and *Bam*HI and cloned into M13.

PCR fragments were labeled with [α-^{32}P]dCTP using the random priming method. The λEMBL3 genomic library prepared with DNA from the Y strain of *T. cruzi* was a gift from Dr. Antonio Gonzalez (Instituto de Parasitología y Biomedicina "Lopez-Neyra", Granada) [10]. Approximately 100000 plaques were screened for the DHFR-TS gene, and positive isolates were plaque-purified. Fragments were subcloned into pUC18 and M13 vectors for DNA sequencing. Plasmid DNA containing inserts were used to transform *E. coli* HB101 cells. Plasmids were purified on CsCl/ethidium bromide gradients [9].

Contour clamped homogeneous electric field electrophoresis. Low-melting-point agarose blocks of *T. cruzi* were prepared as described [11]. Chromosomes were separated on a 1.5% agarose gel in TBE 0.1 M/ 0.5 μg ml^{-1} ethidium bromide using a CHEF system (Pharmacia). The following parameters were used: 40 h at 170 V with a pulse time of 240 s. The resulting gel was transferred to a Hybond-N (Amersham) nylon filter and subjected to Southern blot analysis.

Sequencing of genomic DNA coding for DHFR-TS. The subcloned fragments were sequenced by the dideoxy chain termination method (Kit for DNA sequencing with Sequenase, USB) using pUC forward and reverse universal sequencing primers and other synthetic oligonucleotides. Sequencing gels contained 5% acrylamide and 41% urea in TBE 1 × buffer. Both DNA strands for the coding and untranslated regions were completely sequenced.

Expression system construction. For expression in *E. coli*, the entire coding sequence was amplified by PCR. Oligonucleotide primers for amplification of the DHFR-TS coding sequence (ON3 and ON4) were designed so that *Eco*RI and *Hin*dIII restriction sites were introduced at the 5′ and 3′ ends for convenient cloning in the expression vector. A band of the correct size was purified using the GeneClean Kit (BIO101) and cloned in pKK223.3 (Pharmacia). The expression construct (pKTcTD) was originally cloned and propagated in HB101. Double strand DNA was sequenced by the dideoxy method to confirm the correct sequence after amplification. Subsequently pKTcTD was used to transform the expression host JM105 (lacI^{Q}), the *E. coli* K12 Thy^{-} strain χ2913 and the Thy^{-} DHFR^{-} strain PA414. In the case of χ2913 and PA414, transformation mixtures were plated on duplicate minimal agar plates [12], containing 50 μg ml^{-1} ampicillin, with or without 50 μg ml^{-1} thymidine to test for the Thy^{-} phenotype.

Purification of DHFR-TS and enzyme assays. Bacterial clones were grown in LB containing 50 μg ml^{-1} ampicillin. When isopropyl-1-thio-β-D-galactoside induction was performed, bacterial cells transformed with pKTcTD were first grown to an *A* of 0.5–0.7 at 37°C and then 5 mM isopropyl-1-thio-β-D-galactoside was added; cultures were grown for an additional 3 h, cells were collected by centrifugation and when not used immediately, frozen at −80°C. Recombinant DHFR-TS from *T. cruzi* was purified essentially by affinity chromatography on methotrexate-Sepharose as described for the bifunctional DHFR-TS from *L. major* [5]. Enzyme was eluted with dihydrofolate, pooled and dihydrofolate removed by filtration through a Sephadex G-25 disposable column (PD-10, Pharmacia) before concentration to 5–6 mg ml^{-1} (Centricon, Amicon). Protein content was determined by the method of Bradford using bovine serum albumin as standard [13]. Samples were analyzed on SDS-PAGE slab gels.

DHFR and TS activity determinations were performed spectrophotometrically as described for the *Leishmania* enzyme [7]. One unit of enzyme activity is the amount of enzyme required to produce 1 nmol of product min^{-1} at 25°C.

3. Results

Genomic DNA cloning. The PCR probe was obtained by amplification of a 284-bp fragment of the DHFR-TS gene using as primers degenerate nucleotides complementary to highly conserved regions flanking the dUMP binding site of the TS

domain. Restriction sites were added to the 5′ end of each primer mixture to facilitate cloning of the PCR products using genomic *T. cruzi* DNA as template. Several products were obtained after amplification of *T. cruzi* DNA and aliquots of the PCR reaction were electrophoresed, transferred to nitrocellulose and probed independently with sequences of the *Leishmania major, Trypanosoma brucei* and *Crithidia fasciculata* DHFR-TS genes. We located a PCR product of the expected size (approximately 300 bp) that hybridized to all three probes. The positive fragment was cloned in M13mp18 and the correct identity of the amplified band was confirmed by sequencing. A genomic DNA library constructed in λEMBL3 was screened with the specific PCR probe. We obtained 21 positive clones in the first screening of approximately 100000 plaques. After the third screening, one of six clones (pPRTS-1) was selected for restriction mapping and further characterization. This clone contained a 13.25-kb *Sal*I-*Sal*I insert which possesed the complete coding region of the DHFR-TS gene. A restriction map of the genomic DNA encompassing the DHFR-TS gene is shown in Fig. 1A. *Sal*I digestion of this insert gives three restriction fragments of 2.5, 3.5 and 7.25 kb and *Sal*I-*Eco*RI double digestion gives various fragments. One *Sal*I-*Eco*RI 1.9-kb fragment contains most of the gene minus 100 bp from the amino terminus which is located in the adjacent *Sal*I-*Sal*I 2.5-kb fragment. The *Sal*I-*Eco*RI 1.9-kb and *Sal*I-*Sal*I 2.5-kb portions were cloned in pUC18 giving pTC1.9 and pTC2.5. Different restriction fragments of the coding sequence obtained by digestion of pTC1.9 were inserted in pUC18 and sequenced on both strands using universal and synthetic primers.

Prediction of the initiation codon was based on the comparison of the *T. cruzi* sequence with the DHFR-TS sequence of *L. major*. The deduced amino acid sequence for the coding region of DHFR-TS is shown in Fig. 2. Translation of the open reading frame of 1563 bp yields a polypeptide of 521 amino acids with an estimated molecular mass of 58829. No intervening sequences were present.

Genomic organization of the T. cruzi DHFR-TS gene. Genomic DNA from *T. cruzi* was digested with different endonucleases blotted and probed with a *Sal*I-*Bst*XI fragment which encompasses 1453 bp of the coding region of the DHFR-TS gene (Fig. 1B). Lanes A, C and D correspond to digestions with endonucleases that present unique restriction sites within the coding region. Lane B corresponds to a digestion with *Eco*RI (site outside the coding region). Double digestions with *Eco*RI and either *Sac*I or *Sal*I are shown in lanes E and F, respectively. In all cases there was hybridization to only one or two bands which suggests that the gene is of single copy.

T. cruzi chromosomes were separated by CHEF (Fig. 3A) and probed with the *Sal*I-*Bst*XI fragment. In this case the probe hybridized to two bands of approximately 0.99 and 1.04 Mb (Fig. 3B).

Expression of recombinant DHFR-TS.

The oligonucleotides ON3 and ON4 were designed to use as primers for amplifying the entire coding sequence by PCR in order to construct an expression system. *Eco*RI and *Hin*dIII restriction sites were placed at the 5′ and 3′ ends for cloning in the expression vector pKK223.3, which contains a ribosomal binding site adequately positioned upstream of the *Eco*RI site for efficient translation in *E. coli*. Amplification conditions were as described in the methods section and template DNA consisted of 400 ng of pPRTS-1 which encompasses the entire coding region. A single band of approximately 1.5 kb was visualized after the amplification reaction. The band was excised, cloned in pKK223.3 to give pKTcTD and used to transform HB101 cells. The authenticity of the construct was verified by restriction analysis and sequencing. Plasmid DNA was used to transform different host cells which were grown under conditions suitable for induction of expression and then analysed for the presence of expression products (Table 1). When pTcTD was used to transform *E. coli* JM105 ($lacI^Q$), enzyme activities were determined after induction of expression with IPTG. Transformation of the *E. coli* Thy^- strain χ2913 and the Thy^- and $DHFR^-$ strain PA414 with this construct complemented TS deficiency and expression was further assessed by activity determina-

tions in crude extracts and SDS-PAGE analysis of cell protein. Fig. 4 shows an SDS-PAGE analysis of the soluble expression products in χ2913 cells. As shown in lane A there is a clear band of approximately 59 kDa, similar in size to *L. major* DHFR-TS. Expression in each strain was analyzed by densitometry, showing that DHFR-TS represents about 1–2%, except in the PA414 strain where it accumulates to less than 1%. In all cases DHFR activity was high, and values varied between 150 U mg^{-1} in the PA414 strain and 500 U mg^{-1} in the JM105 strain. TS activity in crude extracts could only be measured adequately when using χ2913 as expression organism. No detectable activity was obtained with JM105 and PA414. Inclusion bodies containing a protein of molecular weight similar to DHFR-TS were detected in χ2913 *E. coli* extracts. We have estimated that *T. cruzi* DHFR-TS represents about 3–4% of total cell protein.

Purification and characterization of T. cruzi DHFR-TS.

T. cruzi DHFR-TS from 2913 *E. coli* extracts

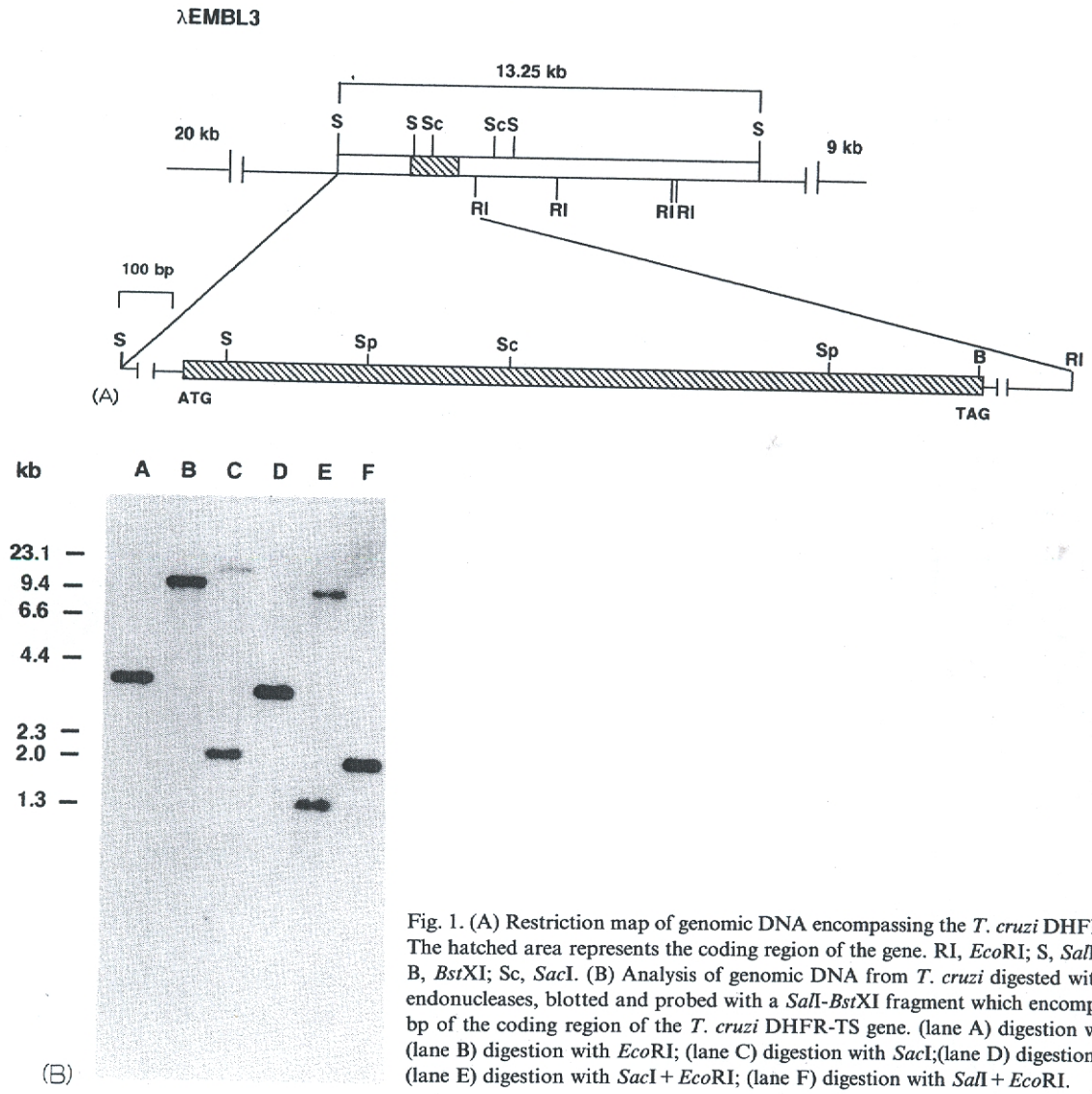
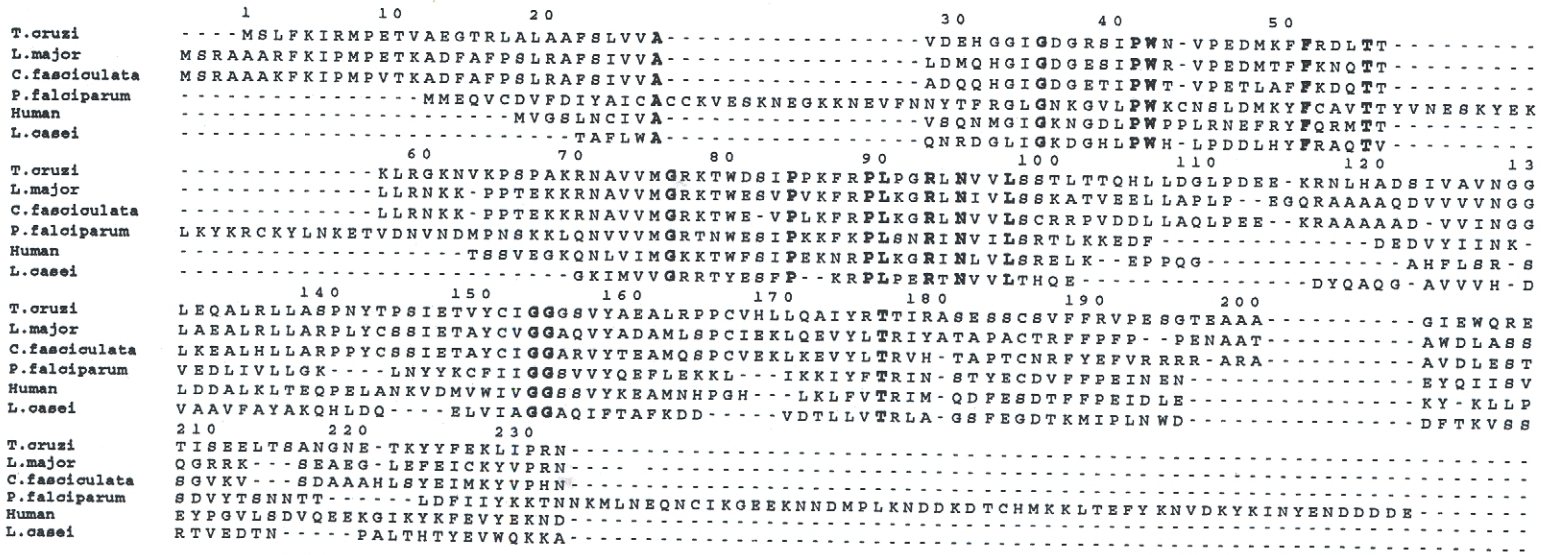


Fig. 1. (A) Restriction map of genomic DNA encompassing the *T. cruzi* DHFR-TS gene. The hatched area represents the coding region of the gene. RI, *Eco*RI; S, *Sal*I; Sp, *Sph*I; B, *Bst*XI; Sc, *Sac*I. (B) Analysis of genomic DNA from *T. cruzi* digested with different endonucleases, blotted and probed with a *Sal*I-*Bst*XI fragment which encompasses 1453 bp of the coding region of the *T. cruzi* DHFR-TS gene. (lane A) digestion with *Bst*XI; (lane B) digestion with *Eco*RI; (lane C) digestion with *Sac*I;(lane D) digestion with *Sal*I; (lane E) digestion with *Sac*I + *Eco*RI; (lane F) digestion with *Sal*I + *Eco*RI.

Dihydrofolate Reductase



Thymidylate Synthase

Fig. 2. Alignment of *T. cruzi* DHFR-TS with selected sequences from other sources. Sequences from *L. major* [20, 35]; *C. fasciculata* [36]; *P. falciparum* [37]; Human (38, 39, 40) and *L. casei* [41,42] are aligned. Residues that are invariant are in bold face.

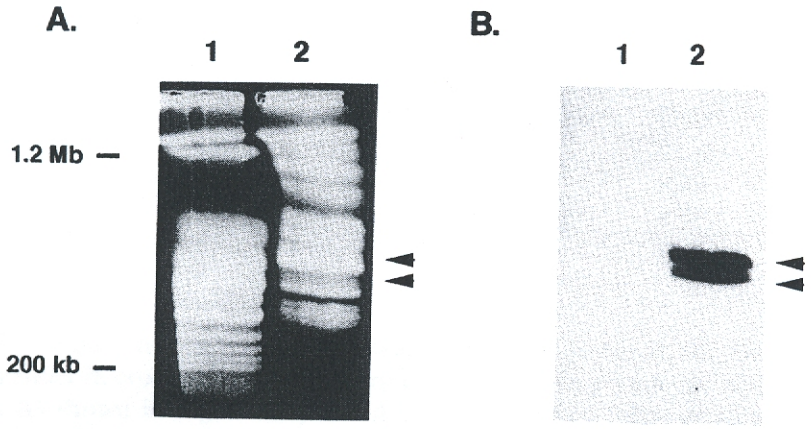


Fig. 3. Contour clamped homogeneous electric field (CHEF) electrophoresis from *T. cruzi*. *A* Ethidium bromide-stained gel. (lane 1) corresponds to *S. cerevisieae* as molecular weight marker chromosomes; (lane 2) corresponds to *T. cruzi* chromosomes. *B* Autoradiograph of a nylon transfer of the gel shown in A probed with the ^{32}P-labeled *SalI-BstXI* fragment.

was purified to homogeneity by MTX-Sepharose affinity chromatography. The conditions used were basically those described for purification of the *Leishmania* bifunctional enzyme [5,7], although 500 mM KCl was necessary in the elution buffer for adequate recovery of protein. The bifunctional protein is quite stable and both enzyme activities are preserved during the purification procedure. Lyophilization and prolonged storage under vacuum at room temperature has no effect on enzyme activity. The purified product gave a single 59-kDa band on SDS-PAGE gels that was indistinguishable from that of *L. major* DHFR-TS. We obtained about 3 mg of apparently homogeneous DHFR-TS from a 2 l culture of χ2913 cells transformed with pKTcTD. Based on DHFR units, we obtain a 41% recovery of protein after purification (Table 2). A comparison of the DHFR/TS ratio in crude extracts with that of the purified enzyme indicates that inactivation does not take place during the purification procedure. DHFR specific activity of the purified protein was 34000 U mg^{-1}, while TS specific activity was approximately 1700 U mg^{-1}. Unlike *L. major*

Table 1
Expression of *T. cruzi* DHFR-TS

Organism	Plasmid	% protein[a]	Specific activity (U mg^{-1})		DHFR/TS
			DHFR	TS	
JM105	pKTcTD	1–2	~500		
PA414	pKTcTD	<1	150–200	–	
χ2913	pKTcTD	1–2	350–495	15–20	~21

[a]The percentage of protein was determined by densitometry analysis and corresponds to soluble cell extract.

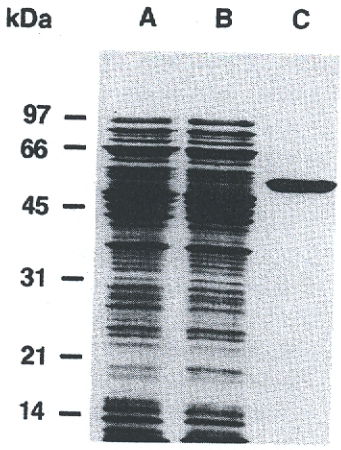


Fig. 4. Purification of recombinant DHFR-TS expressed in *E. coli* strain χ2913. 10% SDS-PAGE stained with Coomassie R250. (lane A) 20 μg crude soluble extract; (lane B) flow-through from methotrexate-Sepharose column chromatography; (lane C) 1 μg of purified DHFR-TS.

Table 2
Purification of *T. cruzi* DHFR-TS

Purification step	Total protein[a] (mg)	DHFR		TS		Purification (-fold)	Yield[b]
		Specific activity (U mg^{-1})	Total activity (U)	Specific activity (U mg^{-1})	Total activity (U)		
Crude extract	700	350	245,000	16	11,200	–	–
MTX-Sepharose	~3	33,500	100,500	1,700	5,100	95	~41%

[a]From 2 l culture.
[b]Based on DHFR activity.

DHFR-TS [5], we have not detected TS degradation products either in crude extracts or in purified preparations. As shown in Fig. 4 the purified bifunctional protein displays a single band on SDS-PAGE gels with an apparent size of 59 kDa. Mass spectrometry analysis of the purified protein gave a molecular mass of 58820 ± 30, a result which is in excellent agreement with the value predicted from the amino acid sequence (58829).

The kinetic parameters for the *T. cruzi* enzyme were determined in experiments where the non-varied substrate was kept at a fixed saturating concentration. The apparent K_m values for DHFR were 17 ± 2 μM and 1.2 ± 0.1 μM for NADPH and H_2folate, respectively. For TS, K_m values were 1.02 ± 0.05 μM for dUMP and 58 ± 5 μM for CH_2-H_4folate. The k_{cat} value obtained for DHFR (72 ± 8 s^{-1}) is considerably higher that was has been described for the recombinant *Leishmania* enzyme expressed in yeast [5] while the turnover number for TS (3.4 ± 0.5 s^{-1}) was comparable to the values reported for the *Leishmania* protein purified from an *E. coli* expression system [5]. There is a considerable variability between the different k_{cat} values determined for *Leishmania* DHFR-TS depending on the source of enzyme and expression system used although most of these differences may be due to differences in N-terminal blocking, proteolytic degradation or instability of the TS domain [7].

4. Discussion

In the present paper we have reported the gene cloning, overexpression in an heterologous system and purification of *T. cruzi* DHFR-TS. As in other protozoa, the DHFR-TS gene in *T. cruzi* is organized with the DHFR domain at the amino terminus, followed by a junction peptide and then the TS domain. The alignment of *T. cruzi* DHFR-TS with other bifunctional and monofunctional counterparts was based on considerations of structural equivalence and maximal sequence similarity [14,15] (Fig. 2). We have assigned amino acids 1–234 (Asn-231 in the *Leishmania* sequence aligns with Asn-234 in *T. cruzi*) to the DHFR domain and the remaining residues have been included in the TS portion. This is supported by the observation that the DHFR domains of *T. cruzi* (unpublished results) and *Plasmodium falciparum* [16] can be expressed independently in a catalytically active form in an expression vector containing the coding sequence for amino acids 1–234.

The DHFR domain shows significant sequence homology with other DHFRs although it has been established that protozoan sequences are more homologous to eukaryotic forms of DHFR than to prokaryotic ones. Secondary structure features in other DHFRs are also conserved in the *T. cruzi* enzyme except for the two β-strands βG and βH located in the carboxyl-terminal region which in *T. cruzi* correspond to the highly variable region defined as the junctional sequence. As in other protozoa, the amino terminus is approximately 20 amino acids longer than in monofunctional DHFRs. These first amino acids constitute a variable region and significant amino acid homology does not begin until approximately residue number 23. A secondary structure prediction [17] for this portion of the sequence assigns a β-strand disposition to the first six amino acids followed by an α-helix that immediately precedes the βA strand found in all forms of DHFR. In addition, the 5′

end of the DHFR domain is 4 amino acids shorter than the *Leishmania* sequence. We have determined the position of the initiating ATG by comparison with other protozoan sequences although 3′ downstream there are several other candidates in the proximity; however, these are ruled out as initiation codons by the mass spectral analysis which shows the protein to be precisely the predicted size. Upstream of the assigned initiation codon there are no methionines in the 324 bp we have sequenced. Comparison of the DHFR amino acid sequence revealed a 50.4% identity with *L. major* (118 identical amino acids); 46.2% with *C. fasciculata* (108 identical amino acids); 23.5% identity with *P. falciparum* (55 identical amino acids); 24.4% with the human enzyme (57 identical amino acids) and 14.1% with *L. casei* (33 identical amino acids). Invariant amino acids involved in catalysis or substrate binding are especially conserved. Thus, of 15 residues described to be involved in dihydrofolate binding in *L. casei* DHFR [18], 8 are identical in *T. cruzi* and at least 5 others are conserved. Of 25 residues that are assigned to NADPH binding in the *L. casei* crystal structure [19], 11 are identical, 8 have conservative changes and 6 are varied. Residues 193 to 234 are what is considered as the junction peptide [20]. No significant homology can be described for this portion of the sequence although of special interest with regard to other protozoan sequences is a 3 amino acid insertion at position 215 in the *T. cruzi* sequence. Secondary structure analysis predicts that residues 209 to 234 form a random coil which is immediately followed by the conserved α-A helical structure of the TS domain.

We have assigned 287 amino acids (residues 235–521) to the TS domain. The TS sequence is highly conserved and all of the conserved residues involved in catalysis and substrate binding identified in other TSs are present in the *T. cruzi* enzyme. Overall homology analysis also reveals that the enzyme is more closely related to the eukaryotic vertebrate sequences than to prokaryotes. Amino acid identity is 78.4% with *L. major* (225 identical amino acids); 78.0% with *C. fasciculata* (224 identical amino acids); 50.9% with *P. falciparum* (146 identical amino acids); 62.4% with human (179 identical amino acids) and 45.6% with *L. casei* (131 identical amino acids). Homology with the human enzyme is higher than that with the *Plasmodium* counterpart. Amino acids 283–290 correspond to a consensus sequence TTKR(K)-X_{2-3}-R(K) described for the folate binding site found in all TSs sequenced to date [20]. The sequence MALPPCH in position 398–404 also corresponds to a consensus sequence obtained for the dUMP binding site (MAL(V)P(A/T)PCH(V)). Sixteen invariant residues of the 25 that appear to line the surface of the cavity that contains the active site [21] have identities in the *T. cruzi* enzyme. Overall amino acid identity of the bifunctional enzyme with that of the trypanosomatid *L. major* is 65.8%.

The DHFR domain has a k_{cat} value of 72 s^{-1}, approximately two-fold higher than what has been described for the enzyme purified from a DHFR-TS overproducing strain of *L. major* [7] and 2.7 fold higher than recombinant DHFR-TS expressed in *E. coli* [5]. The k_{cat} and K_m values for TS are similar to what has been described for the *Leishmania* enzyme [5,7] while DHFR K_m values for both H_2folate and NADPH are somewhat higher.

It has been difficult to definitely establish the ploidy of *Leishmania* and *Trypanosoma*. Analysis of total DNA [22–24] and specific loci in mutants or natural populations have suggested that the genome is diploid [25–28], and Iovannisci and Beverley [29] have indicated that chromosomes of *L. major* bearing housekeeping genes are diploid. Likewise, molecular analyses have shown that the DHFR-TS gene from *L. major* is diploid [25]. In the case of *T. cruzi*, a study on the chromosomal localization of genes that encode for several antigens gives strong evidence for diploidy of the epimastigote stage of the parasite [30].

Moreover, it has been demonstrated that several housekeeping genes are located on chromosomes of different sizes in different isolates of parasites like African trypanosomes [31], *Leishmania* [32], *Plasmodium* [33] and *T. cruzi* [28]. In the present case, the DHFR-TS probe hybridizes to two chromosomes which differ considerably in size. There are two possible explanations for this result: either the double band represents homologous chromosomes that present size polymorph-

ism and *T. cruzi* is at minimum diploid at the DHFR-TS locus or the population is heterogeneous and extensive chromosome rearrangements have occured. The factors contributing to the apparent size variability among *T. cruzi* chromosomes are unkown, but may involve telomere growth and contraction, as observed in *T. brucei* [34], transposable elements or differences in the amount of repetitive sequences. The results obtained with genomic DNA digests using restriction enzymes that cut within or outside the gene suggest that DHFR-TS is present as a single copy in the trypanosomal genome, observation that taken together with the data of chromosomal hybridization suggests that the DHFR-TS gene may be localized on two homologous chromosomes which exhibit size polymorphism and can be resolved by CHEF electrophoresis.

Acknowledgements

These studies were supported by grants from the Spanish Programa Nacional de Investigación y Desarrollo Farmaceticos (FAR91–0427), the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseses (TDR)(ID No.920155 L30/181/83), Plan Andaluz de Investigación (Cod. 3277) and U.S. Public Health Service Research R01 AI 19358 (to DVS). P.R. is a Predoctoral Fellow of the Plan Andaluz de Investigación y Caja General de Ahorros de Granada. R.A. and A.O. are Predoctoral Fellows of the Spanish PFPI of the Ministerio de Educación y Ciencia. Mass spectral data was provided by the UCSF Mass Spectrometry Facility (A.L. Burlingame, Director) supported by the Biomedical Research Technology Program of the National Center for Research Resources, NIH NCRR BRTP 01614. We thank Dr. Antonio Gonzalez for comments on this manuscript.

References

[1] Santi, D.V. and Danenberg, P.V. (1984) Folates in pyrimidine nucleotide biosynthesis. In: Folates and Pteridines, Vol. 1, Chemistry and Biochemistry of Folates (Blakey, R.L. and Benkovic, S.J. eds.), pp. 345–398. John Wiley and Sons, New York.

[2] Ferone, R. and Roland, S. (1980) Dihydrofolate reductase-thymidylate synthetase, a bifunctional polypeptide from *Crithidia fasciculata*. Proc. Natl. Acad. Sci. USA 77, 5802–5806.

[3] Garret, C.E., Coderre, J.A., Meek, T.D., Garvey, E.P., Claman, D.M., Beverley, S.M. and Santi, D.V. (1984) A bifuntional thymidylate synthetase-dihydrofolate reductase in protozoa. Mol. Biochem. Parasitol. 11, 257–265.

[4] Sirawaraporn, W., Sertsrivanich, R., Booth, R.G., Hasnch, C., Neal, R.A. and Santi, D.V. (1988) Selective inhibition of *Leishmania* growth by 5-benzyl-2,4-diaminopyrimidines. Mol. Biochem. Parasitol. 31, 79–86.

[5] Grumont, R., Sirawaraporn, W. and Santi, D.V. (1988) Heterologous expression of the bifunctional thymidylate synthase-dihydrofolate reductase from *Leishmania major*. Biochemistry 27, 3776–3784.

[6] Sirawaraporn, W., Sirawaraporn, R., Cowman, A.F., Yuthavong, Y. and Santi, D.V. (1990) Heterologous expression of active thymidylate synthase-dihydrofolate reductase from *Plasmodium falciparum*. Biochemistry 29, 10779–10785.

[7] Meek, T.D., Garvey, E.P. and Santi, D.V. (1985) Purification and characterization of the bifunctional thymidylate synthase-dihydrofolate reductase from methotrexate-resistant *Leishmania tropica*. Biochemistry 24, 678–686.

[8] Garvey, E.P., Coderre, J.A. and Santi, D.V. (1985) Selection and properties of *Leishmania tropica* resistant to 10-propargyl-5,8-didezafolate, an inhibitor of thymidylate synthetase. Mol. Biochem. Parasitol. 17, 79–91.

[9] Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G.,Smith, J.A. and Struhl, K. (eds) (1987) In *Current Protocols in Molecular Biology*, John Wiley and Sons, New York.

[10] Gonzalez, A., Rosales, J.L., Ley, V. and Diaz, C. (1990) Cloning and characterization of a gene coding for a protein (KAP) associated with the kinetoplast of epimastigotes and amastigotes of *Trypanosoma cruzi*. Mol. Biochem. Parasitol. 40, 233–244.

[11] Garvey, E.P. and Santi, D.V. (1986) Stable amplified DNA in drug-resistant *Leishmania* exists as extrachromosomal circles. Science 233, 535–540.

[12] Belfort, M., Moelleken, A., Maley, G.F. and Maley, F. (1983) Purification and properties of T4 phage thymidylate synthetase produced by the cloned gene in an amplification vector. J. Biol. Chem. 258, 2045–2051.

[13] Read, S.M. and Northcote, D.H. (1981) Minimization of variation in the response to different proteins of the Coomassie blue G dye-binding assay for protein. Anal. Biochem. 116, 53–64.

[14] Blakey, R.L. (1984) In: Folates and Pteridines, Vol. 1, Chemistry and Biochemistry of Folates (Blakey, R.L. and Benkovic, S.J. eds.), pp. 191–253. John Wiley and Sons, New York.

[15] Volz, K.W., Matthews, D.A., Alden, R.A., Freer, S.T.,

Hansch, C., Kaufman, B.T. and Kraut, J. (1982) Crystal structure of avian dihydrofolate reductase containg phenyltriazine and NADPH. J. Biol. Chem. 257, 2528–2536.

[16] Sirawaraporn, W., Prapunwattana, P., Sirawaraporn, R., Yuthavong, Y. and Santi, D.V. (1993) The dihydrofolate reductase domain of *Plasmodium falciparum* thymidylate synthase-dihydrofolate reductase. J. Biol. Chem. 268, 21637–21644.

[17] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. 120, 97–120.

[18] Bolin, J.T., Filman, D.J., Matthews, D.A., Hamlin, R.C. and Kraut, J. (1982) Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. J. Biol. Chem. 257, 13650–13662.

[19] Filman, D.J., Bolin, J.T., Matthews, D.A. and Kraut, J. (1982) Crystal structures of *Escherichia coli* and *Lactobacillus casei* dihydrofolate reductase refined at 1.7 Å resolution. J. Biol. Chem. 257, 13663–13672.

[20] Grumont, R., Washtien, W.L., Caput, D. and Santi, D.V. (1986) Bifunctional thymidylate synthase-dihydrofolate reductase from *Leishmania tropica*: Sequence homology with the corresponding monofunctional proteins. Proc. Natl. Acad. Sci. USA 83, 5387–5391.

[21] Hardy, l. W., Finer-Moore, J.S., Montfort, W.R., Jones, M.O., Santi, D.V. and Stroud, R.M. (1987) Atomic structure of thymidylate synthase: Target for rational drug design. Science 235, 448–455.

[22] Leon, W., Fouts, D.L. and Manning, J.E. (1978) Sequence arrangement of the 16S and 26S rRNA genes in the pathogenic haemoflagellate *Leishmania donovani*. Nucleic Acids Res. 5, 491–504.

[23] Borst, P., van der Ploeg, M., van Hoek, J.F.M., Tas, J. and James, J. (1982) On the DNA content and ploidy of trypanosomes. Mol. Biochem. Parasitol. 6, 13–23.

[24] Scholler, J.K., Reed, S.G. and Stuart, K. (1986) Molecular karyotype of species and subspecies of *Leishmania*. Mol. Biochem. Parasitol. 20, 279–293.

[25] Beverley, S.M., Ellenberger, T.E., Iovannisci, D.M., Kapler, G.M., Petrillo-Peixoto, M.L. and Sina, B.J. (1988) Gene amplification in *Leishmania*. In: Biology of Parasitism, Vol. 9 (Englund, P.T. and Sher, A. eds), pp. 431–438. Liss, New York.

[26] Iovannisci, D.M., Goebel, D., Allen, K., Kaur, K. and Ullman, B. (1984) Genetic analysis of adenine metabolism in *Leishmania donovani* promastigotes: Evidence for diploidy at the adenine phosphoribosyltransferase locus. J. Biol. Chem. 259, 14617–14623.

[27] Gibson, W.C., Osinga, K., Michels, P.A.M. and Borst, P. (1985) Trypanosomes of subgenus *Trypanozoon* are diploid for housekeeping genes. Mol. Biochem. Parasitol. 16, 231–242.

[28] Gibson, W.C. and Miles, M.A. (1986) The karyotype and ploidy of *Trypanosoma cruzi*. EMBO J. 5, 1299–1305.

[29] Iovannisci, D.M. and Beverley, S.M. (1989) Structural alterations of chromosome 2 in *Leishmania major* as evidence for diploidy, including spontaneous amplification of the mini-exon array. Mol. Biochem. Parasitol. 34, 177–188.

[30] Henriksson, J., Åslund, L., Macina, R.A., Franke de Cazzulo, B.M., Cazzulo, J.J., Frasch, A.C.C., and Pettersson, U. (1990) Chromosomal localization of seven cloned antigen genes provides evidence of diploidy and further demostration of kariotype variability in *Trypanosoma cruzi*. Mol. Biochem. Parasitol. 42, 213–224.

[31] Gibson, W.C. and Borst, P. (1986) Size-fractionation of the small chromosomes of *Trypanozoon* and *Nannomonas* trypanosomes by pulsed field gradient gel electrophoresis. Mol. Biochem. Parasitol. 18, 127–140.

[32] Bishop, R.P. and Miles, M.A. (1987) Chromosome size polymorphisms in *Leishmania donovani*. Mol. Biochem. Parasitol. 24, 263–272.

[33] Sharkey, A., Langsley, G., Patarapotikul, J., Mercereau-Puijalon, O., McLean, A.P. and Walliker, D. (1988) Chromosome size variation in the malaria parasite of rodents, *Plasmodium chabaudi*. Mol. Biochem. Parasitol. 28, 47–54.

[34] Bernards, A., Michels, P.A.M., Lincke, C.R. and Borst, P. (1983) Growth of chromosome ends in multiplying trypanosomes. Nature 303, 592–597.

[35] Beverley, S.M., Ellenberger, T.E. and Cordingley, J.S. (1986) Primary structure of the gene encoding the bifunctional dihydrofolate reductase-thymidylate synthase of *Leishmania major*. Proc. Natl. Acad. Sci. USA 83, 2584–2588.

[36] Hughes, D.E., Opeolu, A.,Shonekan, O.A. and Simpson, L. (1989) Structure, genomic organization and transcription of the bifunctional dihydrofolate reductase-thymidylate synthase gene from *Crithidia fasciculata*. Mol. Biochem. Parasitol. 34, 155–166.

[37] Bzik, D.J., Li, W. -B., Horii, T. and Inselburg, J. (1987) Molecular cloning and sequence analysis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene. Proc. Natl. Acad. Sci. USA 84, 8360–8364.

[38] Masters, J.N. and Attardi, G. (1983) The nucleotide sequence of the cDNA coding for the human dihydrofolic acid reductase. Gene 21, 59–63.

[39] Chen, M. -J., Shimada, T., Moulton, A.D., Cline, A., Humphries, R.K., Maizel, J. and Nienhuis, A.W. (1984) The functional human dihydrofolate reductase gene. J. Biol. Chem 259, 3933–3943.

[40] Takeishi, K., Kaneda, S., Ayusawa, D., Shimizu, K., Gotoh, O. and Seno, T. (1985) Nucleotide sequence of a functional cDNA for human thymidylate synthase. Nucleic Acids Res. 13, 2035–2043.

[41] Freisheim, J.H., Bitar, K.G., Reddy, A.V. and Blankenship, D.T. (1978) Dihydrofolate reductase from amethopterin-resistant *Lactobacillus casei*. J. Biol. Chem. 253, 6437–6444.

[42] Maley, G.F., Bellisario, R.L., Guarino, D.V. and Maley, F. (1979) The primary structure of *Lactobacillus casei* thymidylate synthetase. J. Biol. Chem. 254, 1301–1304.